

Deriving the Poisson distribution from the Binomial

Motivation

In this section we show you how the probability model of a binomial model can be translated across to a Poisson process. It teaches us a few things:

- The memoryless property of a binomial process carries across to a Poisson process;
- The Poisson process is often a good approximation to the binomial process; and therefore
- The various distributions of the Poisson process are often good approximations to their corresponding binomial process distributions.

Derivation of the Poisson distribution

We'll start with an example application. Imagine that I am about to drink some water from a large vat, and that randomly distributed in that vat are bacteria. The larger the quantity of water I drink, the more risk I take of consuming bacteria, and the larger the expected number of bacteria I would have consumed.

We could have a go at modeling this as a binomial process. A trial could be a small amount of water, say 1ml. A success would be that there were at least one bacterium in that ml of water. If the concentration in the vat was much less than 1 bacterium/ml, then the probability of having a second bacterium in a contaminated ml is small. Then the number of trials n is the number of ml of water I drink, and the probability of success is roughly the concentration/ml.

I could make the binomial model increasingly more accurate by having smaller units of water in a trial: 0.1ml, 0.01ml, etc. The problem is that for this model to work, I must always test whether the probability of success for a trial is sufficiently low that I have no real chance of a second bacterium in the one unit (that would be bad because the second bacterium wouldn't be accounted for in the 0/1 regime of a binomial process).

What we need to do is make the number of trials approach infinity, so of course the probability of success approaches zero, but keep the same expected level of risk. This is exactly what Simeon Poisson did.

Some mathematics

Consider a binomial process where the number of trials tends to infinity, and the probability of success at the same time tends to zero, with the constraint that the mean of the Binomial distribution = np remains finitely large. The probability mass function of the Binomial distribution is:

$$p(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \quad (1)$$

So, in the example above, x would be the number of bacteria I consume in n units of water, and p is the probability that a random unit of water contains a bacterium.

We'll replace p with the Poisson intensity = bacteria/ml, and the number of trials n with the amount of water consumed t ml. Note that x and t must have matching units, so that:

$$t = np$$

Putting into Equation 1. gives:

$$p(x) = \frac{n!}{x!(n-x)!} \left(\frac{\lambda t}{n}\right)^x \left(1 - \frac{\lambda t}{n}\right)^{n-x}$$

(2)

$$= \frac{n(n-1) \dots (n-x+1)}{n^x} \frac{(\lambda t)^x}{x!} \frac{(1 - \lambda t/n)^n}{(1 - \lambda t/n)^x}$$

For n large and p small:

$$\left(1 - \frac{\lambda t}{n}\right)^x \approx e^{-\lambda t}, \frac{n(n-1) \dots (n-x+1)}{n^x} \approx 1, \left(1 - \frac{\lambda t}{n}\right)^x \approx 1$$

which simplifies Equation 2. to:

$$p(X = x) \approx \frac{e^{-\lambda t} (\lambda t)^x}{x!}$$

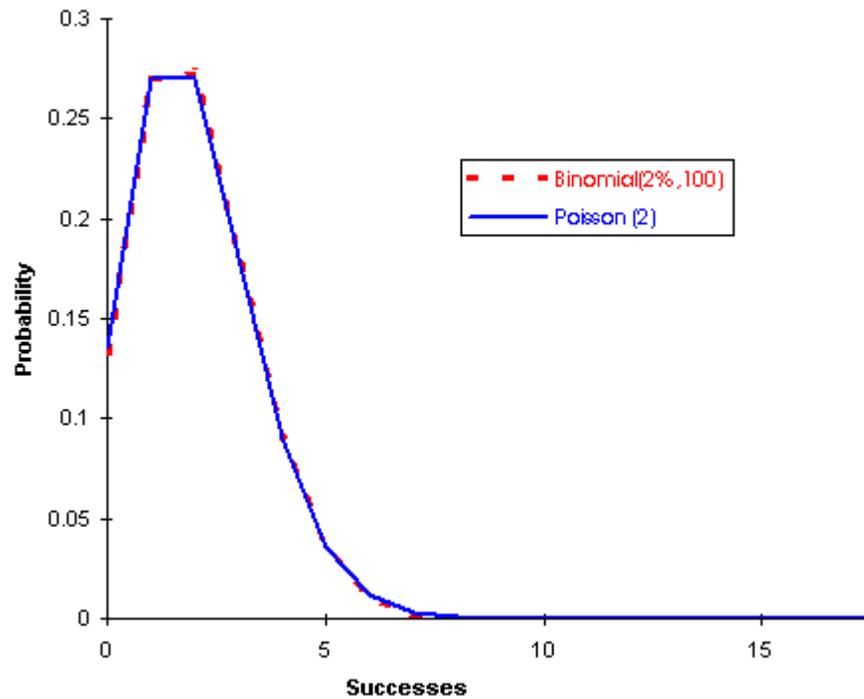
This is the probability mass function for the Poisson(t) distribution.

Number of events in time $t = \text{Poisson}(t)$

when the average number of events that will occur in a unit interval of exposure is known to be .

Poisson approximation to the Binomial

From the above derivation, it is clear that as n approaches infinity, and p approaches zero, a Binomial(p, n) will be approximated by a Poisson($n \cdot p$). What is surprising is just how quickly this happens. The approximation works very well for n values as low as $n = 100$, and p values as high as 0.02. The figure below plots these two distributions together and shows how closely the Binomial distribution is approximated by the [Poisson](#) distribution in such limiting cases.



Poisson distribution is *not* a rare event distribution

The Poisson distribution is often mistakenly considered to be only a distribution of rare events. It is certainly used in this sense to approximate a Binomial distribution, but has far more importance than that, as we've just seen. In a Poisson process, the same random process applies for very small to very large levels of exposure t . So, for example, if cases of sporadic illness occurred in a population at a rate of 100/year, the number of cases occurring in an hour would be Poisson(100/365/24) and the number in the next ten years would be Poisson(100*10): the same process requires the same distribution whether this is likely to be a very low number or very high. The Poisson parameter is λt , which removes any sense of 'time' units, so whether the exposure is 100 years or 1 second is irrelevant: the Poisson distribution is only defined by the expected number of events in that period.
